Articles

The Solvent-Protected Core of the Hairpin Ribozyme−Substrate Complex†

Ken J. Hampel, Nils G. Walter, and John M. Burke*

Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, University of Vermont, Burlington, Vermont 05405

Received May 11, 1998; Revised Manuscript Received August 17, 1998

ABSTRACT: The complex between the hairpin ribozyme and its substrate consists of two domains that must interact in order to form a catalytic complex, yet experimental evidence concerning the points of interaction between the two domains has been lacking. Here, we report the use of hydroxyl radical footprinting to define the interface between the two domains. Cations that support very efficient ribozyme catalysis (magnesium and cobalt(III) hexammine) lead to the formation of a docked complex that features several regions of protection, indicating a solvent-inaccessible core within the tertiary structure of the complex. Cations that are suboptimal in cleavage reactions do not produce complexes with regions of reduced solvent accessibility. Nucleotides encompassing the substrate cleavage site (c−2, a−1, g+1, and u+2) are strongly protected, suggesting their internalization into the catalytic core. Four distinct segments of the ribozyme are protected, including G11−A14, C25−C27, A38, and U42−A43. Protection of these sites is eliminated when g+1, an essential base at the cleavage site, is replaced by A. In addition, mutations which are known to decrease the fraction of docked complexes decrease or eliminate formation of a solvent-inaccessible core. Taken together, these observations demonstrate that we have identified the catalytic core of the active hairpin ribozyme−substrate complex.

RNA molecules that serve important structural or catalytic roles in nature fold into complex three-dimensional structures. To a large extent, our understanding of the biological function of RNA is directly linked to our grasp of structure and folding pathways. Ribozymes are a class of RNAs which can catalyze chemical reactions in the absence of protein cofactors (*1*, *2*). As such, ribozymes provide an ideal opportunity to relate the structure of RNA to its function.

The hairpin ribozyme is a catalytic RNA motif that is derived from the negative-strand genomic RNA of the tobacco ringspot virus satellite (−)sTRSV[1] (*3*, *4*). It catalyzes a sequence-specific, reversible transesterification reaction that results in the endonucleolytic cleavage of an RNA substrate with formation of products containing 5′-OH and 2′,3′-cyclic phosphate termini or in the ligation of RNA molecules with such termini (*3*). Other small ribozymes, including the hammerhead, *Neurospora* VS, and hepatitis delta virus ribozymes, generate similar cleavage products (*5*), but it is not yet clear whether they share similar reaction mechanisms. While the secondary structure of the hairpin ribozyme is well understood, its tertiary folding is not well characterized (*6*). Phylogenetic comparisons, mutational analysis, in vitro genetics, and chemical modification protection mapping have established that the ribozyme−substrate complex is composed of two domains (A and B), each containing an internal loop flanked by two short helical elements (Figure 1A; refs *7*−*10*). The scissile bond is located within internal loop A. Mutations at nearly all sites within the internal loops inhibit activity, and it is believed that specific contacts between the two loops must form in order for a catalytic complex to be generated. Such an interaction would be accompanied by a sharp bend around the junction between helices 2 and 3, which is believed to function as a flexible hinge between the two domains (*11*, *12*). The kinetics of folding into this catalytically proficient docked complex have recently been followed using fluorescence resonance energy transfer (FRET) between domains A and B (*13*). In addition, an inactive alternative form of the ribozyme has been characterized which adopts an extended structure as a result of coaxial stacking of helices 2 and 3 (*14*; J. E. Heckman, N. G. Walter, K. J. Hampel, E. K. O'Neill, and J. M. Burke, in preparation). Further evidence for a functional interaction between the two domains comes from the reconstitution of activity following separation of the linkage between the domains (*15*). The reaction rate approaches that of the intact ribozyme at very high RNA concentrations, suggesting that the interaction between the domains is specific but weak.

† Supported by grants from the National Institutes of Health, a postdoctoral fellowship from the Medical Research Council of Canada to K.J.H., and a Feodor Lynen fellowship from the Alexander von Humboldt Foundation to N.G.W.

* Corresponding author. Phone: (802) 656-8503. Fax: (802) 656-8749. E-mail: jburke@zoo.uvm.edu.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; VS, Varkud satellite; (−)sTRSV, negative strand satellite RNA of tobacco ringspot virus; FRET, fluorescence resonance energy transfer; NMR, nuclear magnetic resonance; HPLC, high-pressure liquid chromatography; ·OH, hydroxyl radical.

10.1021/bi981083n CCC: $15.00
Published on Web 09/30/1998

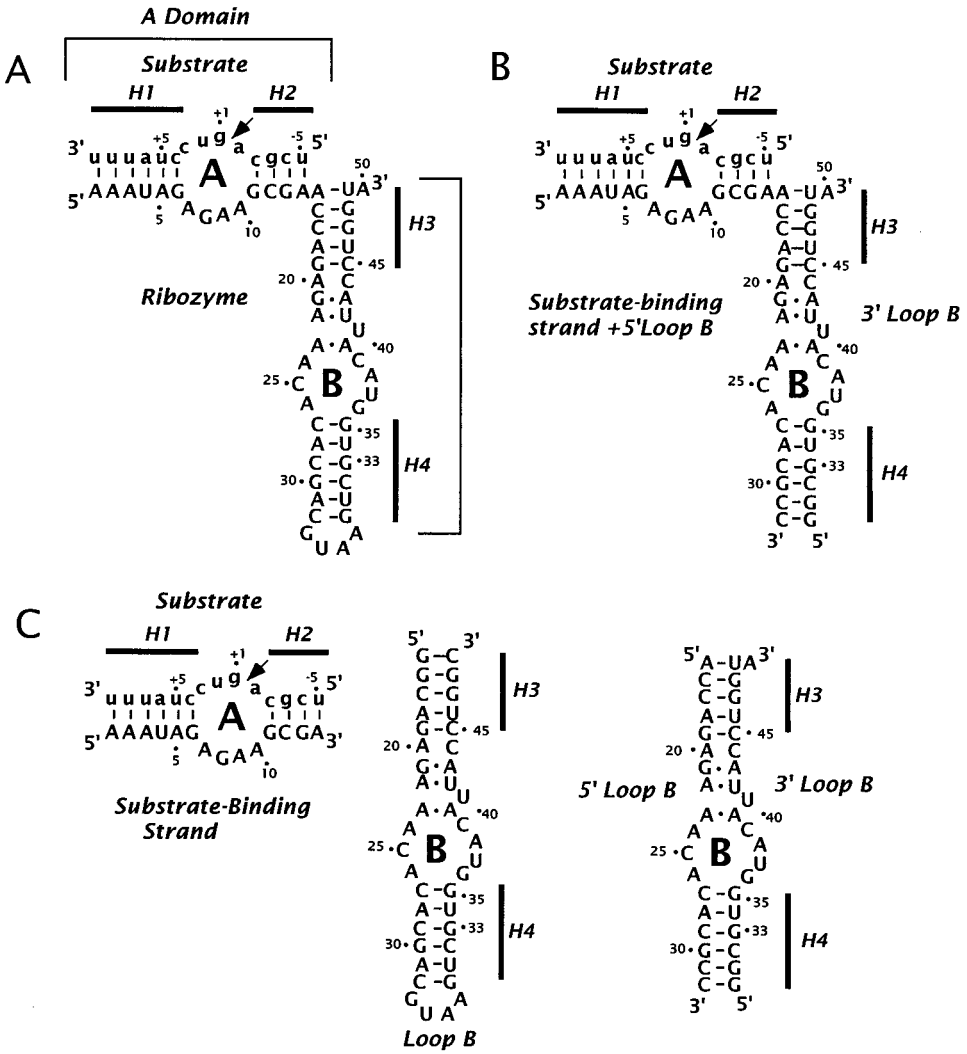

FIGURE 1: Trans-acting hairpin ribozyme constructs used in this study. Ribozymes are derived from the wild-type trans-acting ribozyme with the following modifications. The substrate-binding domain has been optimized to reduce misfolding (*15*, *22*). A U39C mutation has been included which has been shown to increase the activity of the ribozyme (*8*). Finally, helix 4 has been increased from 3 to six base pairs, and the terminal loop has been changed from a GUU to a GUAA loop (*15*, *22*). Numbering is consistent with that of the wild-type trans-acting ribozyme from positions 1−30 and 33−50 (*31*). Between positions 30 and 33 the numbering is from 30.1 to 30.9. (A) One-piece ribozyme−substrate complex. (B) Two-piece ribozyme−substrate complex. (C) Ribozyme components and substrate used to study interdomain interactions from separated loops A and B. The A domain is composed of the internal loop A flanked by helices 1 and 2 (H1 and H2). The B domain is composed of the internal loop A flanked by helices 3 and 4 (H3 and H4).

While little information is available concerning the structure of the docked ribozyme−substrate complex, significant progress has been made toward elucidating the structures of the individual domains. An NMR model of domain A rationalizes some of the known chemical accessibilities and sequence requirements for this domain and suggests a specific pattern of noncanonical base pairing (*16*). In addition, a partial structure for domain B has been deduced. UV cross-linking of loop B pointed to both sequence similarity and structural homology among a family of UV cross-linkable motifs (*17*). The structures of two members of this group, loop E of eukaryotic 5S rRNA and the sarcin−ricin loop of 28S rRNA, have been solved by NMR spectroscopy (*18*, *19*) and are essentially identical.

Recently, published results on the hairpin ribozyme have been used along with new cross-linking data to propose a three-dimensional molecular model of the ribozyme−substrate complex (*20*). In this model, domains A and B are thought to interact through a ribose-zipper motif composed of four 2′-hydroxyl groups (residues 10, 11, 24, and 25) that were previously shown to be important for activity (*21*). This interaction predicts a primary constraint for the register of the interaction between the two domains, with a direct juxtaposition of the lower half of loop B and the nucleotides surrounding the scissile bond. This assembly requires a significant distortion of the connection between domains A and B resulting in unpairing of the A14·u−5 base pair, a notion supported by chemical modification results (*10*).

While previous chemical modification experiments have provided important information about the structure of the individual domains, they shed little light on interdomain interactions, since footprints that could be attributed to interdomain contacts were not obtained (*10*). Although the study provided valuable information about the most stable local interactions within the ribozyme−substrate complex, it used molecules that are now known to display some structural heterogeneity (J. E. Heckman and J. M. Burke, unpublished observations). Consequently, weak interactions, such as those between domains A and B, could not be

detected. Recently, our laboratory has developed sequence variants that result in molecules with greater activity and a lessened propensity for forming alternative, inactive structures (*15*, *22*, *23*).

In the present study, we employ uniformly folding ribozyme−substrate complexes, together with cations known to support efficient ribozyme reactions, to analyze the structure of the docked ribozyme−substrate complex and the requirements for docking of the two domains. To differentiate between the inside and outside of a folded RNA, the solvent accessibility of the C4′ position of ribose can be monitored by addition of an Fe(II)−EDTA complex together with H_2O_2 to the RNA in solution (*24*). The metal−chelator complex generates hydroxyl radicals, which attack the C4′ position of the sugar and result in sugar decomposition and phosphodiester cleavage (*25*); this information can be used to analyze RNA tertiary structure and RNA−protein interactions (*26*, *27*). The high degree of accuracy of this method for predicting solvent accessibilities of RNAs has recently been demonstrated by Cate et al. (*28*).

Here, we report the examination of the hydroxyl radical protected interface between the two domains of the hairpin ribozyme. We observe a solvent-inaccessible core which is largely confined to the segments of the internal loops of the ribozyme−substrate complex. These data provide an experimental tool with which to study the requirements for docking of the two domains and define the register of the interdomain interaction.

MATERIALS AND METHODS

Preparation of Nucleic Acids. All synthetic RNAs and DNAs were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using standard phosphoramidite chemistry. Reagents were obtained from Glen Research. Synthetic RNA was deprotected by the method of Sproat et al. (29) and purified from truncated RNAs by gel electrophoresis using 20% polyacrylamide and 8 M urea gels, followed by removal from incompletely deprotected material by C8 reverse-phase HPLC. The full-length ("one-piece") ribozyme was transcribed using a partial double-stranded DNA template and purified by polyacrylamide and 8 M urea gel electrophoresis as described (*30*, *31*). RNA was 5′-end-labeled with [γ-^{32}P]ATP with T4 polynucleotide kinase following dephosphorylation with calf intestinal alkaline phosphatase.

Hydroxyl Radical Footprinting. Fe(II)−EDTA cleavage reactions were carried out by the method of Tullius and Dombroski (*32*), with slight modifications. Final volumes for all reactions were 10 μL. The 5′-end-labeled RNAs (30 nM), and unlabeled RNAs at a range of concentrations listed in each figure, were added together with 20 mM Tris-HCl (pH 7.5) and cations at the concentrations listed in each figure. The results of experiments employing $Co(NH_3)_6^{3+}$ as a folding cation, and controls without added multivalent cation, were identical when these samples also contained 100 mM NaCl. In this way we were able to ensure that secondary structures assembled as depicted in Figure 1. This mixture was preincubated at 37 °C for 20 min, as suggested by FRET experiments for docking kinetics of the two domains (*13*), and was further tested using the hydroxyl radical footprinting technique. Longer folding times did not result in altered patterns of hydroxyl radical accessibility. Cleavage reaction components, 1 μL of 0.4% v/v H_2O_2, 1 μL of 50 mM sodium ascorbate, and 1 μL of Fe(II)−EDTA [10 mM $Fe(NH_4)_2(SO_4)_2$ and 12 mM Na_2−EDTA, pH 8, mixed immediately prior to the start of the reactions], were added sequentially to the inside wall of the reaction tubes. The use of a 1:2 ratio of Fe(II):EDTA gave identical results. The reactions were initiated by a quick microcentrifuge spin and were allowed to proceed for 3 min in the dark at room temperature. Reactions were terminated by addition of 10 mM thiourea, 90% formamide, 0.02% bromophenol blue, and 0.02% xylene cyanol, and products were separated by denaturing gel electrophoresis. Results were analyzed by autoradiography and by quantitation using a Bio-Rad GS-525 radioanalytic imaging system. Quantification of the accessibility of specific residues, and profiling specific lanes, was carried out by subtracting the background of unreacted material and correcting for gel loading differences. Identification of protected areas by these measurements was very consistent. In general, any site which is protected from ·OH by tertiary folding by greater than 35% was identified as such by phosphorimager analysis. The level of protection between individual experiments varied within 20%. The patterns of protection did not vary between individual experiments.

Ribozyme Kinetics. A ribozyme component mixture (10 μM substrate-binding strand and loop B at concentrations indicated in Figure 6) and 5′-end-labeled substrate were preincubated separately in reaction buffer [20 mM Tris-HCl, pH 7.5, and 1.5 mM $Co(NH_3)_6^{3+}$] at 37 °C for 20 min. Reactions were initiated by addition of the substrate mix to the ribozyme components and incubated at 37 °C. Eight time points were taken over 30 min (for domain B concentrations < 1 μM) or 2 h (for domain B concentrations ≥ 1 μM) by quenching aliquots of the reaction into 3 volumes of 8 M urea, 0.02% bromophenol blue, and 0.02% xylene cyanol on ice. Data were collected as described above. Rate constants were derived from the initial linear phase of each reaction, and the average of three experiments was plotted as a function of loop B concentration. The average error in these experiments was less than 10%.

Curve Fitting. The dependence of rate constants and protection amplitudes on loop B or cation concentrations (for example, Mg^{2+}) were fitted to the cooperative binding equation

$$f = f_{max} \frac{[Mg^{2+}]^n}{[Mg^{2+}]^n + (K^{Mg}_D)^n} \quad (1)$$

to yield an apparent dissociation constant K_D for Mg^{2+}. All fits were calculated with Microcal Origin 4.1 software employing Marquardt−Levenberg nonlinear least-squares regression.

RESULTS

Hydroxyl Radical Footprinting Reveals an Internalized Core of Active Hairpin Ribozyme−Substrate Complexes. The RNA molecules in Figure 1 were used to determine if nucleotides of the ribozyme become protected from attack by hydroxyl radicals (·OH) upon folding of the ribozyme−substrate complex. Their catalytic properties have been described previously, and it was found that their folding is

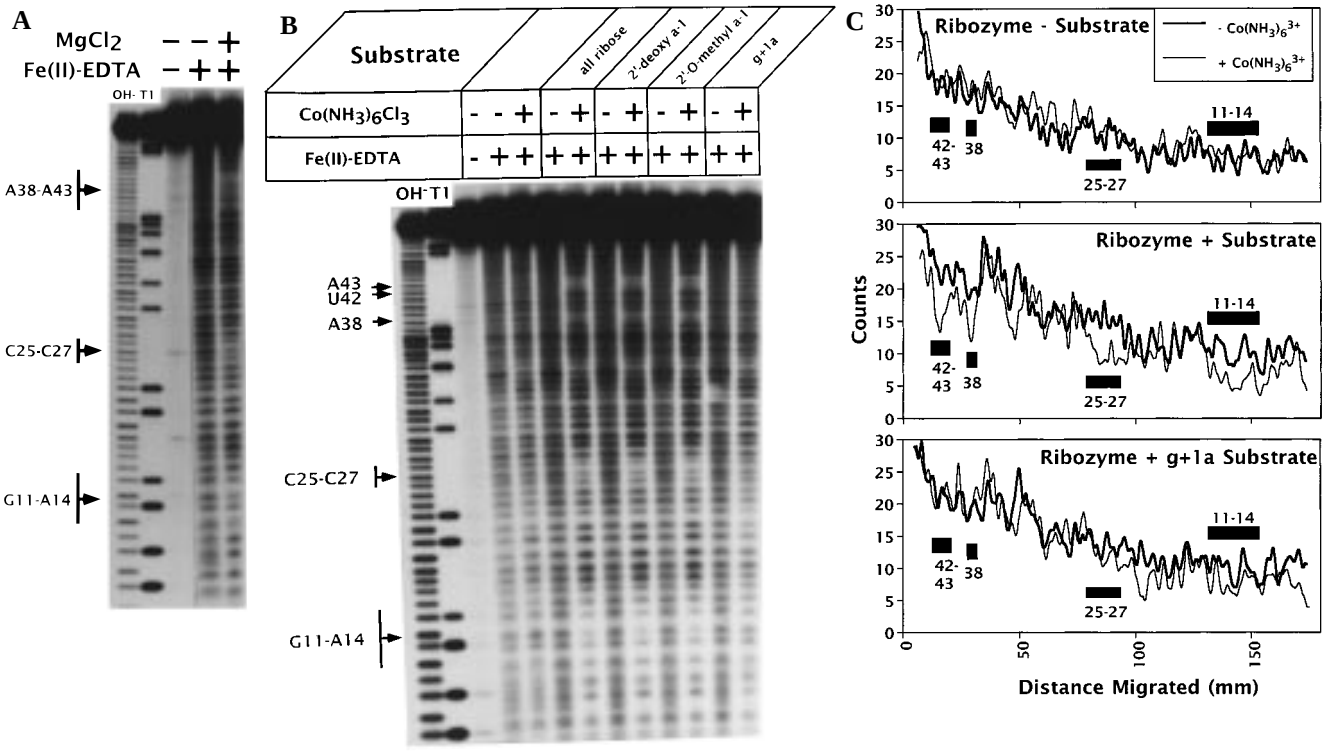

FIGURE 2: Formation of a solvent-inaccessible core in the tertiary structure of the ribozyme−substrate complex. The 5′-end-labeled one-piece ribozyme (30 nM) was incubated with 0.2 μM substrate (exception: all-ribose substrate, 2 μM) and 20 mM Tris-HCl, pH 7.5, for 10 min at 37 °C in the presence or absence of 1.5 mM $Co(NH_3)_6^{3+}$, or 12 mM Mg^{2+}, as indicated. Fe(II)−EDTA reactions were carried out at room temperature for 3 min, and the reactions were stopped and loaded onto a 10% polyacrylamide and 8 M urea gel. Arrows on the left side of each panel indicate protected residues. Note that the identity of each band on the Fe(II)−EDTA ladders is actually one band below the corresponding band on partial ribonuclease T1 digest and alkali hydrolysis ladders. This is due to complete removal of the nucleotide by the hydroxyl radical attack (*25*). (A) Polyacrylamide gel analysis of a ribozyme−substrate complex reacted with Fe(II)−EDTA reagent in the presence or absence of Mg^{2+}. (B) Polyacrylamide gel analysis of a ribozyme reacted with Fe(II)−EDTA reagent in the presence or absence of different substrates and $Co(NH_3)_6^{3+}$. (C) Phosphorimager quantification of gel analysis. Fe(II)−EDTA-mediated cleavage of the ribozyme was quantified and is represented as follows: Cleavage (counts) was plotted as a function of the distance migrated for each RNA species. The sequence was inferred from comparison with partial ribonuclease T1 digest and alkali hydrolysis ladders (not shown). Thin and thick lines represent ribozyme in the presence and absence of 1.5 mM $Co(NH_3)_6^{3+}$, respectively. Black bars represent regions of the ribozyme which are consistently protected from hydroxyl radical cleavage in the presence of all-ribose substrate or substrates that have been modified at the 2′ position of the a−1 residue.

considerably more homogeneous than that of the (−)sTRSV "wild-type" ribozyme (*15*, *22*).

Protection of the ribozyme was assayed using 5′-^{32}P-labeled ribozyme in the presence or absence of different metal ions and substrate (Figure 2). Mg^{2+} ions support a robust reaction at concentrations of approximately 10 mM in vitro and are believed to be important in supporting catalysis in the cellular environment. However, recent results show that very similar levels of activity are obtained in the presence of Mg^{2+} or relatively low concentrations ($\leq$1 mM) of cobalt(III) hexammine (*33*, *35*). This indicates that inner-sphere coordination of components of the RNA chain is not required for catalysis.

When Mg^{2+} was added to ribozyme plus substrate, three regions of protection were observed, indicating that the C4′ positions of the protected nucleotides have become less accessible to solvent (Figure 2A). The protected sites include the ribozyme segment of helix 2 (G11−A14) and two sites within internal loop B, C25−C27 and A38−A43. In the presence of $Co(NH_3)_6^{3+}$ (Figure 2B,C), the pattern of protection of ribozyme plus substrate is identical except C39−U41 is not protected. We have consistently observed that $Co(NH_3)_6^{3+}$ produces a slightly stronger and cleaner protection pattern than that obtained in the presence of Mg^{2+}. Similar results were obtained when the $Co(NH_3)_6^{3+}$ buffer solution contained 100 mM NaCl, in agreement with data suggesting that ribozyme activity in the presence $Co(NH_3)_6^{3+}$ is slightly enhanced by the addition of NaCl (*33*). Protection patterns were also examined in the presence of cations that support ribozyme cleavage reactions less efficiently than $Co(NH_3)_6^{3+}$ or Mg^{2+}, or not at all. Co^{2+}, Ca^{2+}, Mn^{2+}, and Sr^{2+} at 50 mM, 20 mM spermidine, and 300 mM Na^+ did not promote the formation of an observable, solvent-inaccessible core structure. Thus, the appearance of an internalized core structure depends on the presence of cations that optimize ribozyme activity.

The pH dependence of this folding transition was investigated since the lack of a strong pH dependence for catalysis (*34*), and domain docking (*13*), has been documented for the hairpin ribozyme and for this construct in particular (A. R. Banerjee, J. A. Esteban, and J. M. Burke, manuscript in preparation). Over the pH range 6.2−9.0, in the presence of $Co(NH_3)_6^{3+}$, the pattern and the amplitude of ribozyme protections are identical (data not shown). This result also rules out the possibility that the results shown in Figure 2 are buffer-specific. Since the extrinsic requirements for the observed folding transition are identical to those required for cleavage, and because cleavage could be observed during

the course of the footprinting experiment (data not shown), it is likely that the observed protection pattern results from a tertiary structure that represents an authentic intermediate in the catalytic cycle.

Substitution of Catalytically Critical Residues Prevents Domain Docking. Two classes of noncleavable substrate modifications have been previously identified which permit binding to the ribozyme with affinities roughly similar to those of the unmodified substrate, as demonstrated by inhibition, gel mobility-shift, and fluorescence-quenching studies (*23*, *31*, *36*, *37*). The first class involves replacement of the 2′-hydroxyl group of a−1, which serves as the attacking nucleophile in the cleavage reaction, with a 2′-*O*-methyl group or with a hydrogen atom (*23*, *31*, *37*). The second class contains substitutions of the guanosine immediately to the 3′ side of the scissile bond (g+1) with A, C, or U (*31*). No cleavage of either class of inhibitor has been observed; given our limits of resolution, we estimate that the cleavage rate is decreased by a minimum of 5 orders of magnitude.

Hydroxyl radical protection experiments were conducted using both classes of inhibitors, and dramatic differences were identified (Figure 2B,C). Protection with the analogues containing a 2′-deoxy or 2′-*O*-methyl sugar at the cleavage site showed the same pattern as observed for the unmodified substrates. Compared to the cleavable substrate, significantly lower concentrations of the noncleavable analogues were required to confer maximal protection, presumably because substrate cleavage and product dissociation that occur during the assay necessitate a large molar excess of cleavable substrate.

In contrast to the results using the 2′-substituted inhibitors, no protection was obtained with the noncleavable g+1a analogue (Figure 2B,C). Therefore, although the two classes of inhibitors prevent cleavage of bound substrate, they act through two different mechanisms. Substituting g+1 with other bases prevents association of the two domains, while modification of the 2′-OH blocks a step following docking, presumably reaction chemistry.

The Reactive Phosphodiester Linkage Is Internalized in the Docked Complex. To evaluate the solvent accessibility of the substrate in the folded complex, we carried out Fe(II)−EDTA reactions using radiolabeled substrate (Figure 3). Several RNA constructs were used, and folding was assayed in the presence or absence of ribozyme and under a variety of ionic conditions. In the presence of both excess cold ribozyme and Mg^{2+}, substrate nucleotides spanning the cleavage site (c−2, a−1, g+1, and u+2) became strongly protected. Identical results were obtained in the presence of $Co(NH_3)_6^{3+}$, but not with any of the other cations tested (data not shown). Protection of the substrate was observed only when both the substrate-binding strand and domain B are present (Figure 3 and data not shown). A substrate with a g+1a substitution was also end-labeled and tested for ribozyme-dependent folding (data not shown). Although a variety of conditions have been tested, we have never observed protection of substrate that does not contain G at position +1. This result supports the proposal that the g+1a substitution strongly inhibits docking of the A and B domains. The general result that C4′ atoms close to the cleavage site are internalized is important, since binding and cleavage of the substrate RNA by the ribozyme is likely to be accompanied by folding of the ribozyme core very close to the scissile bond in the substrate.

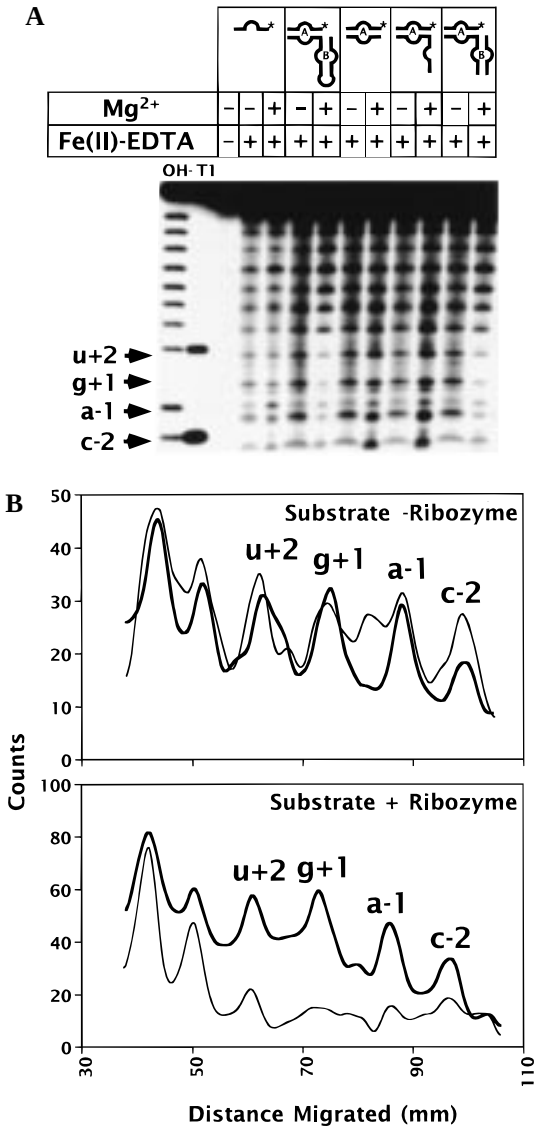

FIGURE 3: The cleavage site is internalized in the ribozyme−substrate complex. (A) Denaturing 20% polyacrylamide gel analysis of 5′-end labeled 2′-deoxy-a−1-substituted substrate reacted with Fe(II)−EDTA reagent in the presence or absence of different cold ribozyme components (1 μM each) and 12 mM Mg^{2+}. Diagrams of the reaction components are shown above each set of lanes. Experiments were carried out as described in Figure 2 and in Materials and Methods. (B) Phosphorimager quantification of gel analysis. Thin and thick lines represent ribozyme in the presence and absence of 12 mM Mg^{2+}, respectively. Fe(II)−EDTA-mediated cleavage of the substrate was quantified and is represented as described in Figure 2.

Cleavage site internalization is also sensitive to mutations in the ribozyme (Figure 4). Kinetic and FRET analyses of ribozyme−substrate complexes with A50U and D50 mutations in the ribozyme show that these changes reduce the cleavage rates, and the proportion of docked molecules, with D50 causing the most severe decreases (*13*). The results in Figure 4 show that, in the presence of Mg^{2+}, ·OH protection of the substrate is likewise impaired by these changes. These data strongly support the view that only an active ribozyme−substrate complex is producing the observable solvent-inaccessible core structure.

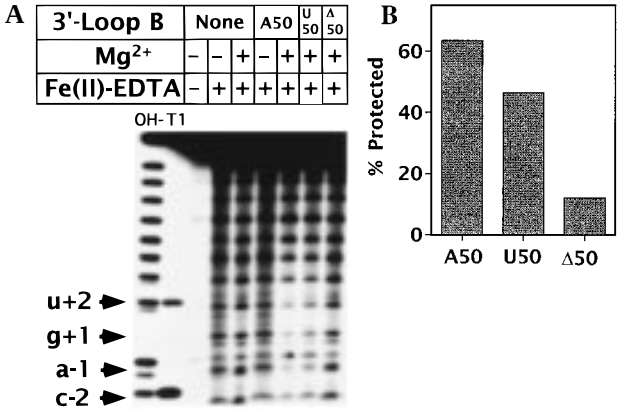

FIGURE 4: Internalization of the cleavage site is sensitive to mutations that disrupt interdomain docking. (A) Denaturing 20% polyacrylamide gel analysis of 5′-end-labeled 2′-deoxy-a-1-substituted substrate reacted with Fe(II)−EDTA reagent in the presence or absence of cold two-piece ribozymes and mutants (1 μM) and 12 mM Mg^{2+}. (B) Phosphorimager quantification of gel analysis. Fe(II)−EDTA-mediated cleavage of the substrate at position g+1 was quantified as described in Materials and Methods. Fraction protected was calculated relative to a control experiment without 3′ loop B RNA added.

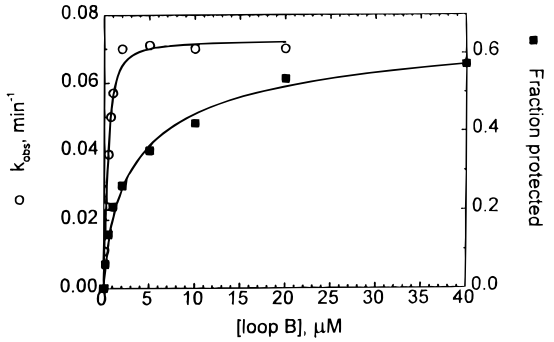

FIGURE 5: Ribozyme cleavage rate (k_{obs}) and hydroxyl radical protection of position g+1 as a function of loop B concentration. The 5′-end-labeled 2′-deoxy-a−1 substrate was incubated with the substrate-binding strand (10 μM) and isolated loop B in the presence of 1.5 mM $Co(NH_3)_6^{3+}$. Kinetics and ·OH radical experiment samples were separated on denaturing 20% polyacrylamide gels. Reactions were carried out, and the results were quantified as described in Materials and Methods. Fraction protected was calculated relative to a control experiment without loop B RNA added.

$Co(NH_3)_6^{3+}$ Supports a Stronger Interaction between the A and B Domains than Mg^{2+}. To test the concept that the interdomain contacts in $Co(NH_3)_6^{3+}$-containing solutions may be stronger than those obtained in Mg^{2+}, we measured the apparent $K_M^{\text{domain B}}$ for a reaction using the ribozyme construct shown in Figure 1C (left panel), in which the two domains are physically separated. When Mg^{2+} is present, the $K_M^{\text{domain B}}$ is very high for this reaction (270 μM; ref *15*). Figure 5 shows the result of the same single-turnover reaction carried out in the presence of $Co(NH_3)_6^{3+}$. The $K_M^{\text{loop B}}$ in these reactions was 0.4 ± 0.1 μM, suggesting that $Co(NH_3)_6^{3+}$ permits a much stronger interaction between the separated domains than does Mg^{2+}. We have also studied the protection of the substrate as a function of loop B concentration by quantifying the ribozyme-dependent protection of g+1 (Figure 5). The K_D that we observed for this interaction is 4.1 ± 1.8 μM.

The cation concentration requirements for core formation also illustrate differences between Mg^{2+} and $Co(NH_3)_6^{3+}$.

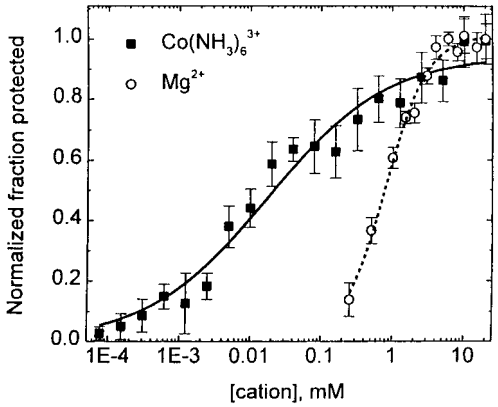

FIGURE 6: $Co(NH_3)_6^{3+}$ is more potent than Mg^{2+} at promoting the formation of the docked ribozyme−substrate complex. Plot of protection from hydroxyl radical cleavage at position g+1 of a ribozyme−substrate complex as a function of Mg^{2+} and $Co(NH_3)_6^{3+}$ concentrations. The 5′-end-labeled 2′-deoxy-a−1 substrate was reacted with Fe(II)−EDTA in the presence of unlabeled ribozyme (0.5 μM) and the indicated cation. Samples were separated on denaturing 20% polyacrylamide gels. Protection of the g+1 position was quantified relative to a negative control in which ribozyme RNA was omitted as described in Materials and Methods. Although the protection pattern observed in these experiments was identical, the level of protection from independent experiments was found to vary somewhat. Thus, the protection values from independent experiments were normalized to the values of the experiment that gave the highest level of protection. Plotted values are the average of three independent experiments.

Protection of the substrate residue g+1 by a one-piece ribozyme was assayed in the presence of varying concentrations of these two cations (Figure 6). The apparent average K_D values for $Co(NH_3)_6^{3+}$ and Mg^{2+} are 17 ± 0.5 μM and 780 ± 50 μM, respectively. Thus, a 50-fold higher concentration of Mg^{2+} is required to achieve the same level of protection afforded by $Co(NH_3)_6^{3+}$. These data are in close agreement with the relative cation concentration requirements for cleavage reactions (*34*) and further support our view that the complex identified here is part of the folding pathway for the active ribozyme. One other feature of the Mg^{2+} and $Co(NH_3)_6^{3+}$ concentration curves is their difference in shape. Significant increases in protection are observed over nearly 4 orders of magnitude of $Co(NH_3)_6^{3+}$ concentration. The Hill coefficients for both cations are below 1; thus cation binding is noncooperative. One possible explanation for this result is that $Co(NH_3)_6^{3+}$ ions may bind to several different sites in the ribozyme, with different apparent K_Ds, and contribute independently to stability of the docked complex.

Domain B Becomes Partially Protected by the Substrate-Binding Strand Alone. Hairpin ribozymes retaining full catalytic activity can be assembled from two (*38*) or three strands (*14*) in which the loop-capping helix 4 has been removed. The 3′ segment of the ribozyme (termed 3′ loop B) was radiolabeled, and protection was examined in a solution containing 1.5 mM $Co(NH_3)_6^{3+}$. In contrast to the results obtained with the full-length ribozyme, protection of A38, U42, and A43 was observed in the absence of substrate (Figure 7, lane 5). This protection is dependent on $Co(NH_3)_6^{3+}$, the 5′ segment of domain B (nucleotides 15−32), and the substrate-binding strand (nucleotides 1−14) (Figure 7) and is not observed in Mg^{2+} (data not shown). The protected sites in the 3′ segment of domain B are identical

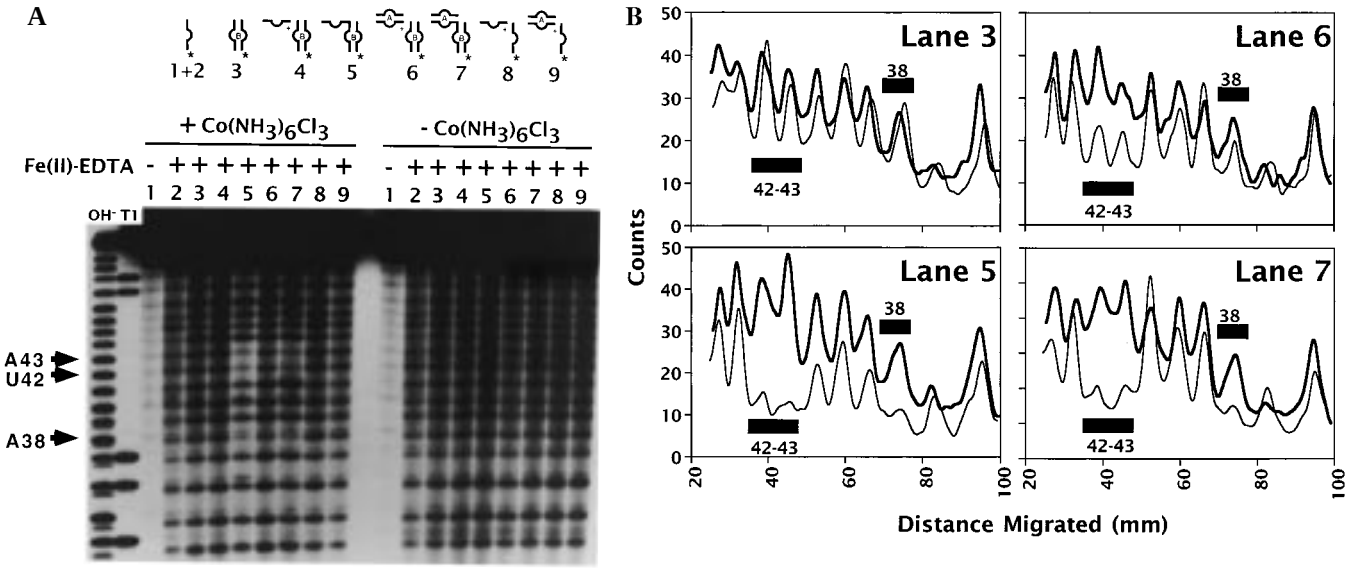

FIGURE 7: $Co(NH_3)_6^{3+}$ promotes the formation of an internalized ribozyme core in the absence of substrate. (A) Denaturing 20% polyacrylamide gel analysis of 5′-end labeled 3′ loop B reacted with Fe(II)−EDTA reagent in the presence or absence of different cold ribozyme components (10 μM each) and 1.5 mM $Co(NH_3)_6^{3+}$. Lane numbers refer to diagrams of the reaction components shown above the gel. Experiments were carried out as described in Figure 2 and Materials and Methods. (B) Phosphorimager quantification of gel analysis. Fe(II)−EDTA-mediated cleavage of the 3′ loop B was quantified and is represented as described in Figure 2.

to those observed when the complete ribozyme−substrate complex is used. No protection of C25−A27, G11−A14, or any other nucleotides within the 5′ segment of domain B has been observed in the absence of substrate or a substrate analogue. While the protection of the 3′ segment of loop B cannot be directly attributed to binding of any specific portion of the substrate-binding strand, the interaction appears to be structurally similar to that of the intact ribozyme−substrate complex.

Protection of these loop B sites was not observed in the full-length ribozyme construct in the absence of substrate or an appropriate substrate analogue. It is conceivable, however, that removal of the hairpin loop which caps helix 4 confers additional flexibility to the complex. Consistent with this hypothesis, substitution of the GUAA tetraloop for the wild-type GUU loop or a UUCG tetraloop does not permit protection of loop B by the substrate-binding strand (data not shown).

In an attempt to further examine the interaction between the substrate-binding strand and domain B, we used a construct in which the two pieces were separated (Figure 1C). No protection of domain B was observed (Figure 7, lane 4), even at very high concentrations of the substrate-binding strand alone (100 μM). When a duplex between the substrate and substrate-binding strand was used, some protection is observed (Figure 7, lane 6). Therefore, the interaction between the substrate-binding strand and domain B appears to be very weak, so that the complex is too short-lived to generate significant protection, unless its components are covalently linked and sterically flexible as in the two-piece ribozyme.

Docking in the Presence of Cleavage Product Analogues Permits the Assignment of Interdomain Contacts. The hairpin ribozyme ligates RNA molecules corresponding to the cleavage products, with a rate greater than 10-fold higher than that of the cleavage reaction (*22*, *39*). However, dissociation rates of the cleavage products are much greater than that of the substrate, so that high concentrations of the cleavage products are required to saturate the ribozyme (*22*, *39*). To achieve adequate concentrations for the mapping experiments, we used synthetic RNA molecules. The RNA corresponding to the 5′ cleavage product contains a 3′-phosphate rather than a 2′,3′-cyclic phosphate. In the presence of $Co(NH_3)_6^{3+}$, but not Mg^{2+}, protection of some but not all of the domain B sites was observed upon the addition of only the 3′ cleavage product (Figure 7). Nucleotides 38 and 42−43 are strongly protected, while nucleotides 11−14 of the substrate-binding strand are partially protected. The additional protected sites within domain B, nucleotides 25−27, were observed upon the subsequent addition of the 5′ cleavage product, but were not observed in the absence of the 5′ cleavage product (Figure 7B). Nucleotides 38 and 42−43 are strongly protected, while nucleotides 11−14 of the substrate-binding strand are partially protected. The additional protected sites within domain B, nucleotides 25−27, were observed upon the subsequent addition of the 5′ cleavage product but were not observed in the absence of the 5′ cleavage product (Figure 7B). No protection of any part of the ribozyme was observed even when a very high concentration (100 μM) of the 5′ product was added alone. Protection of nucleotides 38 and 42−43 may arise from direct interaction between the part of loop A formed by binding of the 3′ product, or the 3′ product permits another region, possibly G11−A14, to make the contact. Protection of C25−C27 requires an intact substrate or binding of both cleavage products. Our interpretation of this result is that the substrate residues that are protected (c−2 through u+2) interact with loop B and protect C25, A26, and C27.

No significant differences could be detected between the protection observed by the substrate or substrate analogue and the protection seen when both cleavage products were present at saturating concentrations. Therefore, the structures of the docked complexes with substrate or cleavage products are similar, at the level of resolution obtained by this method. Similarly, the protection patterns obtained were indistinguishable when the 5′ cleavage product analogue contained

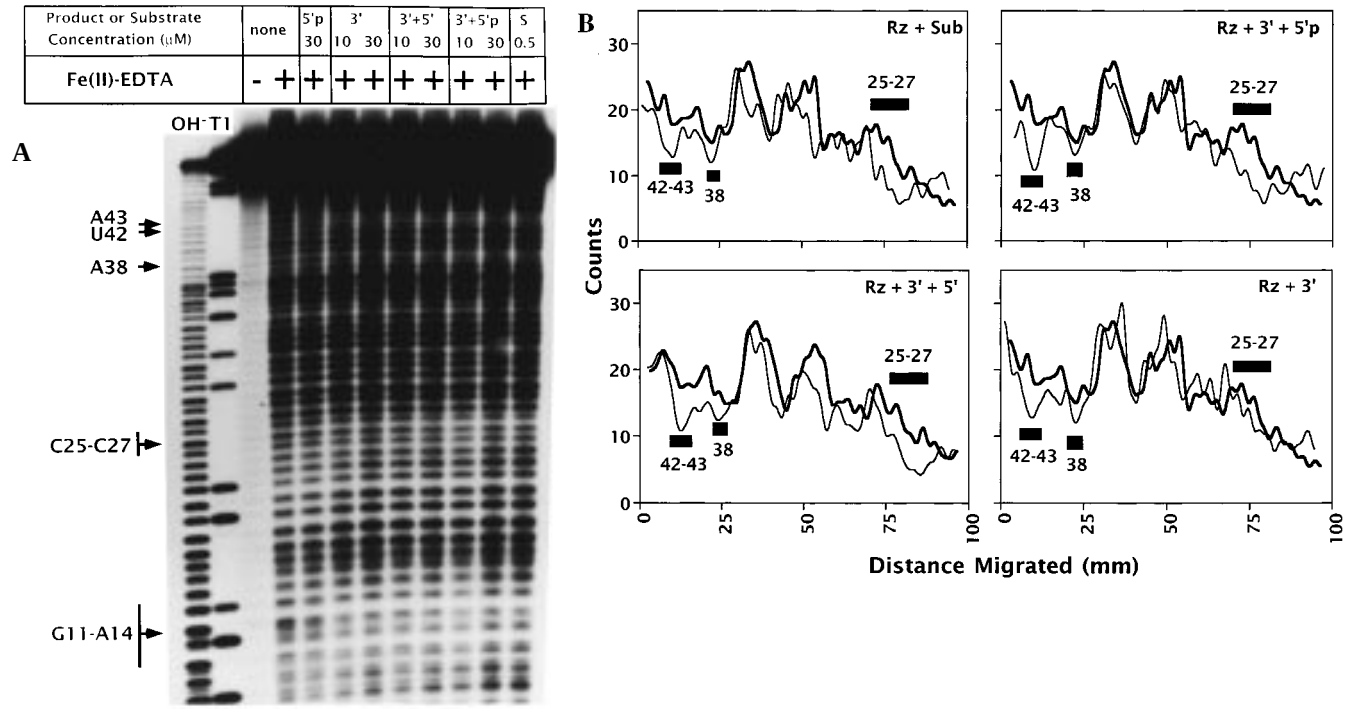

FIGURE 8: Binding of ribozyme cleavage products or 3′ cleavage product leads to the formation of a solvent-inaccessible core. (A) Denaturing 10% polyacrylamide gel analysis of 5′-end-labeled one-piece ribozyme reacted with Fe(II)−EDTA reagent in the presence or absence of different cleavage products or substrate at the concentrations shown and with 1.5 mM $Co(NH_3)_6^{3+}$. Experiments were carried out as described in Figure 2 and Materials and Methods. (B) Phosphorimager quantification of gel analysis. Fe(II)−EDTA-mediated cleavage of the ribozyme was quantified and is represented as described in Figure 2. Abbreviations: 3′, 3′ product; 5′, 5′ product with 3′-OH terminus; 5′p, 5′ product with 3′-phosphate terminus; S, 2′-deoxy-a−1 substrate.

a 3′-phosphate or a 3′-hydroxyl group.

One other notable feature of these experiments is that higher concentrations of the 3′ product were required to obtain protection of residues 38, 42, and 43 to an extent equivalent to that observed when both products are present (Figure 7A and data not shown). In the presence of Mg^{2+}, docking is observed when both of the cleavage products are present, but not when only the 3′ product was presented to the ribozyme (data not shown), providing further support to the notion that $Co(NH_3)_6^{3+}$ facilitates folding of the complex more efficiently than does Mg^{2+}.

DISCUSSION

Although the tertiary structure of the complex between the hairpin ribozyme and its substrate has not yet been determined, it is clear that the complex consists of two domains which fold independently and that an active complex is only formed after the two domains dock to form a closed complex. Recently, our laboratory has devoted considerable effort to begin to understand the requirements for interaction of the two domains and to map the tertiary structure of the complex. The development of variants that fold homogeneously has been an essential prerequisite to the success of this work (*14*, *15*, *22*; J. E. Heckman and J. M. Burke, unpublished observations).

Recently, we have used FRET to measure the kinetics of transitions between the open and closed forms of the ribozyme−substrate complex (*13*). The FRET study showed that docking of the two domains is essential and rate-limiting for the ligation reaction and is essential but is not rate-limiting for cleavage. We demonstrated that many of the mutations and modifications that had previously been shown to eliminate activity act to inhibit a conformational step after substrate binding, by preventing docking of the two domains. Strikingly, the only modifications that have been shown to block cleavage at a step following docking are those in which the attacking nucleophile, the 2′-hydroxyl group at the cleavage site, was replaced by a hydrogen or an *O*-methyl group. However, a significant limitation of the FRET study is that it does not provide any information about the tertiary structure of the closed complex.

The hydroxyl radical footprinting studies described here are consistent with these recent findings. In addition, they considerably extend the results of the fluorescence work, in that they provide the first specific information through which the interface between the two domains can be defined. Previously, this method has been used to identify protected regions in RNA and a group I ribozyme (*24*) and the hepatitis delta ribozyme (*40*). Our results demonstrate that the relatively small hairpin ribozyme is also capable of assuming a tertiary structure containing a solvent-inaccessible region. The fact that the protected region includes the substrate cleavage site together with essential components of the ribozyme suggests that reaction chemistry may take place in an environment where bulk solvent is excluded. Therefore, it is conceivable that the microenvironment of the active site may have properties that are distinct from those exhibited by solvent-accessible RNA, and this may partially account for two unanticipated properties exhibited by this ribozyme, the pH independence of the cleavage and ligation reactions (*34*) and the absence of requirements for divalent metal ions or for inner-sphere coordination of metal ions in catalysis (*33*−*35*).

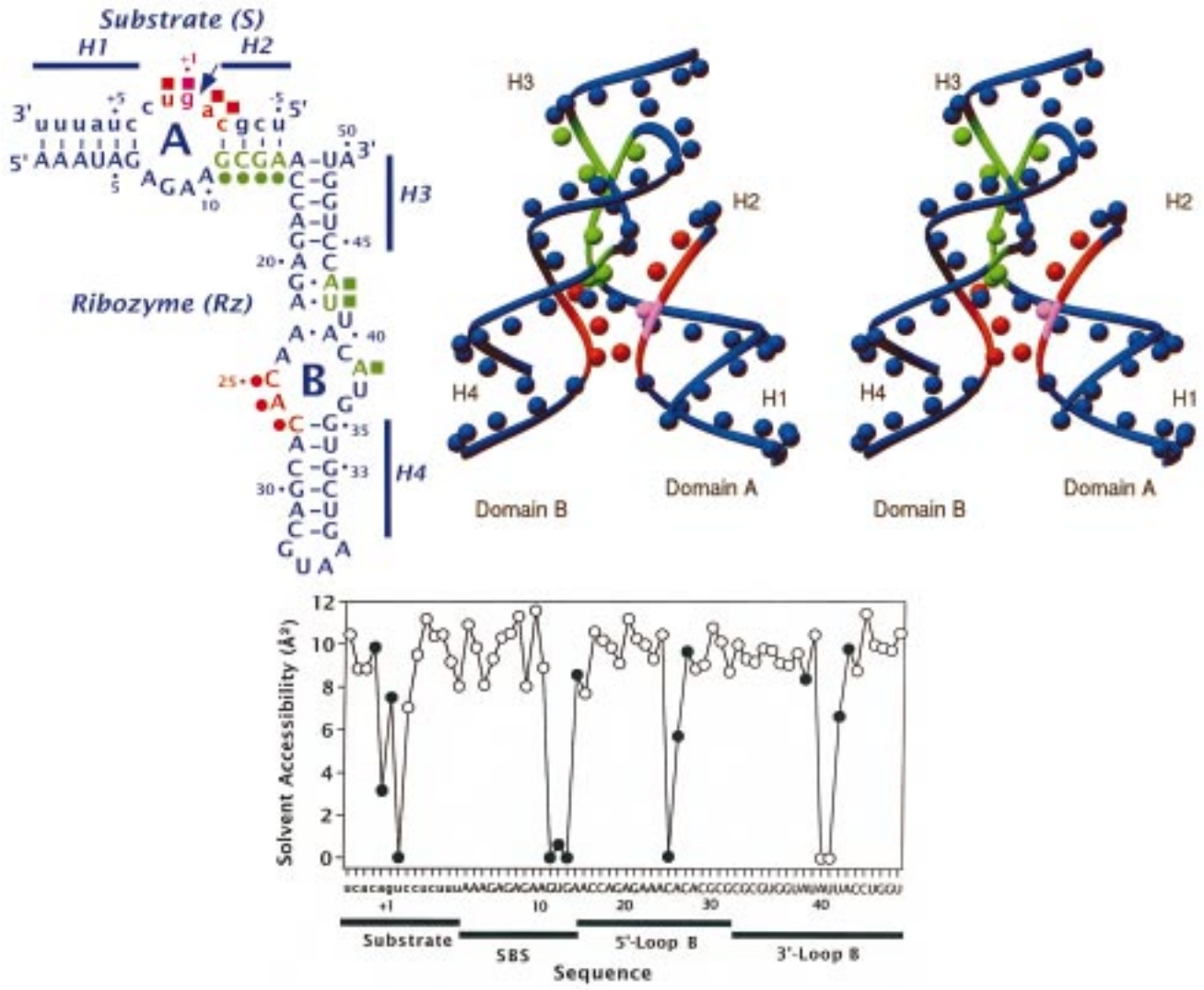

FIGURE 9: Summary of protected sites in the hairpin ribozyme−substrate complex and a comparison with the molecular model for this complex of Earnshaw and co-workers (*20*). (A, top left) The secondary structure of the hairpin ribozyme is shown with squares and circles representing sites of strong and moderate protection, respectively. Color code: blue, unprotected residues in the complex; red and green, pairs of protected areas in the complex which are predicted to interact in the docked complex; pink, critical g+1 residue. (B, top right) Stereoview of the ribbon model of the hairpin ribozyme−substrate complex of Earnshaw et al. (*20*). C4′ atoms are shown as spheres. The color scheme is as indicated in panel A. (C, bottom) Computed solvent accessibility of C4′ positions in the Earnshaw et al. (*20*) model plotted against the sequence. Hydroxyl radical footprinting data are superimposed on computed solvent accessibility by labeling protected and unprotected residues as filled and open symbols, respectively. Calculations of solvent accessibility of the molecular model without hydrogens were made using the solvation module of Insight II software, assuming a solvent radius of 1.4 Å.

What is the evidence that the complex that we have mapped represents a structure that is active and is a conformational intermediate in the reaction pathway? First, ribozyme kinetics carried out with the same RNA constructs and under the conditions used in these studies indicates that greater than 80% of the substrate-bound ribozymes cleave the substrate with the same fast rate (*22*). Furthermore, the inactive substrate-bound ribozyme conformer has been identified as resulting from coaxial stacking of helices 2 and 3, a structure that is unlikely to be capable of rendering specific sites in the complex protected from ·OH (*14*). Our footprinting results suggest that ribozymes bound to substrate are structurally homogeneous. In addition, constructs that give more heterogeneous kinetic populations, for example, the wild-type ribozyme sequence, give very poor hydroxyl radical footprinting results (S. E. Butcher and J. M. Burke, unpublished observations). Theoretically, it is possible that different ribozyme−substrate conformers could be present, with each giving rise to a specific protected site or a unique combination of protections. However, these structurally distinct ribozyme-substrate complexes would have to catalyze the reaction with the same fast rate. This is an unlikely alternative to the hypothesis that a single, stable, docked conformation is responsible for our footprinting results.

Second, the pH dependence and cation requirements for ·OH protection closely parallel those observed for catalysis. However, no protection is seen under ionic conditions that are incompatible with activity. Third, the solvent-inaccessible core of this ribozyme construct is largely confined to the internal loop residues of the complex, which are conserved or contain catalytically critical functional groups, and Watson−Crick base-paired residues, which are strongly conserved as purines or pyrimidines (*8*, *41*−*44*). Our data are, thus, very consistent with the intrinsic requirements for

cleavage activity and the hypothesis that the active folding pathway requires docking of loops A and B.

Fourth, mutations that decrease or eliminate the catalytic activity of the complex similarly affect the footprints which we have observed. The g+1a mutation in the substrate has been previously shown to render the substrate uncleavable and eliminates docking the two domains measured by FRET analysis (*13*, *36*). In our experiments, g+1a mutants do not assemble into complexes where domains A and B are docked. In addition, mutating the A50 position of the ribozyme has specific effects on the •OH footprint which parallel activity and domain docking assays with such mutants. Our primary goal in these experiments was to evaluate the hypothesis that we have mapped the tertiary structure of the active hairpin ribozyme−substrate. Future footprinting experiments on ribozyme−substrate complexes possessing mutations or functional group substitutions will yield additional information about the specific requirements for the formation of a solvent-inaccessible core structure.

What do our results tell us about the folding pathway for the hairpin ribozyme−substrate complex? Identical hydroxyl radical protection patterns are observed when the modification reaction is done during the course of a cleavage reaction containing a saturating concentration of substrate and when a noncleavable substrate analogue that permits docking (as measured by footprinting and FRET) is employed. Since we can show that increasing the folding time of these complexes does not alter the protection patterns, our results demonstrate that some conformational or chemical step, which follows docking, is relatively slow. A very similar conclusion has been reached by FRET analysis of the docking event (*13*). Additional evidence for this idea is the result that $K_D^{\text{loop B}}$ is 10-fold greater than $K_M^{\text{loop B}}$ in experiments where the A and B domains are physically separated (Figure 5). This is an expected result if a conformational step, which occurs after docking of the substrate domain (loop A) and domain B, is rate-limiting for cleavage (*45*).

The protection results also help us to understand the role that metal ions play in folding, since we are now able to monitor both folding and catalysis. It is clear that the docked complex is able to form at much lower concentrations of $Co(NH_3)_6^{3+}$ than Mg^{2+}. This strong correlation between the ability of metal ions to support catalysis and folding is consistent with the hypothesis that the sole function of metal ions in hairpin ribozyme catalysis is to assist in folding the active structure. This hypothesis is supported by recent findings that high concentrations of monovalent cations support efficient RNA cleavage reactions, even in the presence of excess EDTA (J. B. Murray, A. A. Seyhan, N. G. Walter, J. M. Burke, and W. G. Scott, submitted). In contrast, metal ions serve both folding and catalytic roles in the *Tetrahymena* ribozyme. For example, Ca^{2+} permits the correct folding of the ribozyme, but Mg^{2+} or Mn^{2+} is required to support the chemistry of splicing reactions (*26*, *46*, *47*).

Analysis of the protection patterns observed in the presence of cleavage products, individually and in combination, permits us to predict which regions of the two domains are likely to interact with one another. The finding that the protection in the 5′ segment of loop B is only observed when substrate sequences located to both the 5′ and 3′ sides of the cleavage site are present is strong evidence that the cleavage site docks onto the segment of loop B that includes C25, A26, and C27. The other region of protection may involve an interaction between ribozyme sequences comprising helix 2 (G11, C12, G13, and A14) and sequences in the 3′ segment of loop B. We have no direct evidence for this interaction; however, the juxtaposition of the cleavage site with C25, A26, and C27 should place these helix 2 residues adjacent to the protected sites in the 3′-half of loop B. The $Co(NH_3)_6^{3+}$-dependent interaction between the substrate-binding strand and the same sites in loop B supports this hypothesis (Figure 7).

Recently, Earnshaw et al. (*20*) have proposed a tertiary structure model of the complex between the hairpin ribozyme and its substrate. To develop this model, new information from conformational constraint experiments was combined with the body of published structure−function data derived from a number of previous studies. To define the alignment of the two domains, these authors made the plausible, but as yet unproved, assumption that the functionally important 2′ hydroxyl groups at A10 and G11 in domain A and at A24 and C25 (*21*) of domain B interact with one another to form a ribose zipper. Our laboratory has subsequently shown that two of these 2′-OH groups are important for interdomain docking (*13*). To evaluate this model, we used the coordinates to calculate the accessibility of the C4′ positions of all nucleotides in the ribozyme−substrate complex and compared them with the experimentally determined accessibilities obtained in the hydroxyl radical footprinting experiments. Results are shown in Figure 9.

The model predicts four regions where the C4′ positions are expected to have little or no accessibility to solvent. The regions of protection predicted by the model correspond reasonably well with the four protected segments observed in the footprinting experiments (Figure 9). The experimentally observed protections at substrate positions A−1 and U+2 are consistent with the model, as are three of the four sites within helix 2 (G11, C12, and G13) and one of the three sites in the region of loop B that interacts with substrate.

We conclude that the Earnshaw et al. model has accurately predicted the regions of domain A that interact with domain B but that it is less accurate in its prediction of how domain B participates in the interaction. This is not surprising, since domain A has fewer degrees of freedom, and an NMR tertiary structure model is available (*16*). In contrast, no high-resolution data have been obtained for domain B, and essentially nothing is known about the structure of the H4-proximal segment of loop B. The results presented in this paper show that further analysis of this part of the ribozyme is of high importance, since it is the site into which the scissile bond docks and so it may constitute a significant part of the active site of the hairpin ribozyme.

ACKNOWLEDGMENT

We are grateful to Joyce Heckman and Robert Pinard for helpful discussions and for providing some of the RNAs used in this study, to David Pecchia for RNA synthesis, and to Jeff Bond for help with preparing Figure 9.

REFERENCES

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) *Cell 31*, 147−157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983*) Cell 35*, 849−857.

3. Buzayan, J. M., Gerlach, W. L., and Bruening, G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8859−8862.
4. Feldstein, P. A., Buzayan, J. M., and Bruening, G. (1989) *Gene 82*, 53−61.
5. Sigurdsson, S. T., Thomson, J. B., and Eckstein, F. (1998) in *RNA Structure and Function* (Simons, R. W., Ed.) pp 339−376, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
6. Burke, J. M., Butcher, S. E., and Sargueil, B. (1996) *Nucleic Acids Mol. Biol. 10*, 129−143.
7. DeYoung, M. B., Siwkowski, A. M., Lian, Y., and Hampel, A. (1995) *Biochemistry 34*, 15785−15791.
8. Berzal-Herranz, A., Joseph, S., Chowrira, B. M., Butcher, S. E., and Burke, J. M. (1993) *EMBO J. 12*, 2567−2574.
9. Anderson, P., Monforte, J., Tritz, R., Nesbitt, S., Hearst, J., and Hampel, A. (1994) *Nucleic Acids Res. 22*, 1096−1100.
10. Butcher, S. E., and Burke, J. M. (1994) *J. Mol. Biol. 244*, 52−63.
11. Feldstein, P. A., and Bruening, G. (1993) *Nucleic Acids Res. 21*, 1991−1998.
12. Komatsu, Y., Koizumi, M., Nakamura, H., and Ohtsuka, E. (1994*) J. Am. Chem. Soc. 116*, 3692−3696.
13. Walter, N. G., Hampel, K. J., Brown, K. M., and Burke, J. M. (1998) *EMBO J. 17*, 2378−2391.
14. Esteban, J. E., Walter, N. G., Kotzorek, G., Heckman, J. E., and Burke, J. M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
15. Butcher, S. E., Heckman, J. E., and Burke, J. M. (1995) *J. Biol. Chem. 270*, 29648−29651.
16. Cai, Z. P., and Tinoco, I. (1996) *Biochemistry 35*, 6026−6036.
17. Butcher, S. E., and Burke, J. M. (1994) *Biochemistry 33*, 992−999.
18. Wimberly, B., Varani, G., and Tinoco, I. (1993) *Biochemistry 32*, 1078−1087.
19. Szewczak, A. A., Moore, P. B., Chan, Y.-L., and Wool, I. G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9581−9585.
20. Earnshaw, D. J., Masquida, B., Muller, S., Sigurdsson, S. T., Eckstein, F., Westhof, E., and Gait, M. J. (1997) *J. Mol. Biol. 274*, 197−202.
21. Chowrira, B. M., Berzal-Herranz, A., Keller, C. F., and Burke, J. M. (1993) *J. Biol. Chem. 268*, 19458−19462.
22. Esteban, J. E., Banerjee, A. R., and Burke, J. M. (1997) *J. Biol. Chem. 272*, 13629−13639.
23. Walter, N. G., and Burke, J. M. (1997) *RNA 3*, 392−404.
24. Latham, J. A., and Cech, T. R. (1989) *Science 245*, 276−282.
25. Wu, J. C., Kozarich, J. W., and Stubbe, J. (1983) *J. Biol. Chem. 258*, 4694−4697.
26. Celander, D. W., and Cech, T. R. (1991) *Science 251*, 401−407.
27. Weeks, K. M., and Cech, T. R. (1995) *Cell 82*, 221−230.
28. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K. H., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.
29. Sproat, B., Colonna, F., Mullah, B., Tsou, D., Andrus, A., Hampel, A., and Vinayak, R. (1995) *Nucleosides Nucleotides 14*, 255−273.
30. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
31. Chowrira, B. M., and Burke, J. M. (1991) *Biochemistry 30*, 8518−8522.
32. Tullius, T. D., and Dombroski, B. A. (1985) *Science 230*, 679−681.
33. Hampel, A., and Cowan, J. A. (1997) *Chem. Biol. 4*, 513−517.
34. Nesbitt, S., Hegg, L. A., and Fedor, M. J. (1997) *Chem. Biol. 4*, 619−630.
35. Young, K. J., Gill, F., and Grasby, J. A. (1997) *Nucleic Acids Res. 25*, 3760−3766.
36. Chowrira, B. M., Berzal-Herranz, A., and Burke, J. M. (1991) *Nature 354*, 320−322.
37. Walter, N. G., Albinson, E., and Burke, J. M. (1997) *Nucleic Acids Symp. Ser. 36*, 175−177.
38. Chowrira, B. M., and Burke, J. M. (1992) *Nucleic Acids Res. 20*, 2835−2840.
39. Hegg, L. A., and Fedor, M. J. (1995) *Biochemistry 34*, 15813−15828.
40. Rosenstein, S. P., and Been, M. D. (1996) *Biochemistry 35*, 11403−11413.
41. Chowrira, B. M., Berzal-Herranz, A., and Burke, J. M. (1993) *Biochemistry 32*, 1088−1095.
42. Joseph, S., Berzal-Herranz, A., Chowrira, B. M., Butcher, S. E., and Burke, J. M. (1993) *Genes Dev. 7*, 130−138.
43. Grasby, J. A., Mersmann, K., Singh, M., and Gait, M. J. (1995) *Biochemistry 34*, 4068−4076.
44. Schmidt, S., Beigelman, L., Karpeisky, A., Usman, N., Sorensen, U. S., and Gait, M. J. (1996) *Nucleic Acids Res. 24*, 573−581.
45. Fersht, A. (1985) in *Enzyme structure and mechanism*, pp 98−120, W. H. Freeman, New York.
46. Grosshans, C. A., and Cech, T. R. (1989) *Biochemistry 28*, 6888−6894.
47. McConnell, T. S., Herschlag, D., and Cech, T. R. (1997) *Biochemistry 36*, 8293−8303.

BI981083N